Research Data Publication in Principle and Practice

Angus Whyte
Digital Curation Centre

Jonathan Tedds
University of Leicester

Carly Strasser
California Digital Library

Rebecca Lawrence
F1000Research

Sarah Callaghan
Science &Technology Facilities Council

Overview

Data publication is as old as the first academic journals, and its latest incarnations share the goal of communicating research to generate new research cycles from old. New forms of digital research object continue to be innovated to augment the entire research cycle, from pre-proposal Data Management Plan to post-project digital article and underlying datasets. These genres are rapidly evolving, but certain requirements emerge as recurring themes in the effort towards more seamless support for research.

Data repositories can provide raw material for data journals and 'enhanced articles'. The stakeholders include (at least) researchers, publishers, journals, research institutions and funders. Dialogue between them has surfaced the need for new tools, conventions and pooled resources, as envisioned in the FORCE11 Manifesto for example, and further dialogue is essential to progress.

The year since the IDCC'13 data publication workshop on repository accreditation in Amsterdam has seen a number of landmark activities and reports, including: -

- o Reports and workshops from the Jisc funded PREPARDE project (Peer REview for Publication & Accreditation of Research data in the Earth sciences) including multi disciplinary work on peer review, repository accreditation and cross linking for data publishing
- o New data journals and partnerships to link data repositories and journals, e.g. F1000Research, figshare and PLOS , Geoscience Data Journal and British Atmospheric Data Centre
- o Formation of a joint Interest Group on Publishing Data by the international Research Data Alliance and World Data System. Four Working Groups are working on data publishing workflows, bibliometrics, cost models and publishing services
- o Publication of the CODATA-ICSTI Task Group on Data Citation Standards and Practices
- o Reports and work by the Jisc funded PRIME (Publisher, Repository and Institutional Metadata Exchange) project focussed on archaeology and the JORD project which catalogued journal data policies

     o   G8 Science Ministers' statement recommending data publication.

Objectives and outcomes

The full-day IDCC14 data publication workshop will initially provide participants with a general introduction to these and related initiatives. Later sessions and breakout groups will engage participants in addressing opportunities and challenges for data publication, identifying exemplars that can support further progress and collaboration on a number of themes. The themes are likely to include: -

1. Exposing the wide range of data publishing workflows across disciplines and institutions and identifying common challenges, cost implications and solutions
2. Standards and conventions to support citation and attribution, quality assurance and peer review, and metrics for tracking these to ensure high quality data publication is rewarded
3. Improved registry services to broker exchanges of metadata and trust information across repositories, institutions and publishers.
4. Incremental steps that data repositories can take towards repository trust accreditation, to underpin cross-domain sharing and reuse and augment their domain knowledge and reputation.
5. Effective data sharing policy and funding body support measures to improve distribution of expertise across repositories and institutions.
6. Synergies between research institutions' delivery of faculty advice and consultancy services e.g. in data management planning, and support for data publication.

Draft programme

8.45 – 9.00 Arrival
9.00 – 9.15 Aims and Introduction
9.15 – 9.40 Overview of the data publication landscape
9.40 – 10.05 Partnerships and progress towards data publication infrastructure

10.05 – 10.30 tea/coffee

10.30 – 11.20 Data Repository and Institutional perspectives ('the nodes')
Presentation 3
Presentation 4

11.20 – 12.30 Discussion groups #1
Possible themes: models for data peer review; publishing services

12.30 – 13.30 Lunch

13.30 – 14.20 Publisher & service provider perspectives ('the network')
Presentation 5
Presentation 6

14.20 – 15.00 Discussion groups #2 –

Possible themes: identifying workflows, costs, common challenges and solutions; alt-metrics and bibliometrics

15.00 – 15.20 tea/coffee

15.20-15.55 Discussion groups #2 continued

15.55 – 16.30 Report back, discussion & next steps

Speakers to be confirmed.

Intended audience: the workshop will interest anyone involved in: publishing and reviewing datasets, trustworthy digital repository accreditation, data standardisation or policy development.